DOCUMENT RESUME

ED 045 150 LI 002 367

AUTHOR	Webb, Linda; Turner, James
TITLE	Searching and Editing Bibliographic Records Using the IBM Administrative Terminal System (ATS).
INSTITUTION	Five Associated Univ. Libraries, Syracuse, N.Y.
REPORT NO	FTM-70-1
PUB DATE	May 70
NOTE	13p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$0.75
DESCRIPTORS	Automation, *Bibliographic Coupling, *Information Retrieval, *Information Storage, *Information Systems, Libraries, *Search Strategies
IDENTIFIERS	*Administrative Terminal System, ATS

ABSTRACT

IBM's Administrative Terminal System (ATS) is an effective data collection device, but its searching capabilities are severely limited. Searching an ATS document line-by-line is a time consuming process requiring slow printout speeds. A search can operate only upon 9,999 lines of text at any one time. Working storage limitations make it necessary to maintain a quantity of separate ATS documents in permanent storage; each of which is brought in as needed and then returned to storage, so that a series of duplicate searches is necessary if the total file exceeds 9,999 lines. Storage costs as well as access limitations indicate that ATS work is practical only with relatively small and highly fluid files, unless it is used purely for data capture and text editing functions. Costs can be reduced by putting data into off-line storage when not needed and later returning it to the system, but this is not very practical in the present system configuration. To be accessible, data must be maintained in a rigid ATS format. Random searching cannot be done except in a very limited sense. In general, this project was considered worthwhile as a demonstration of the potential and limitations of terminal systems such as ATS. (MF)

ED045150

A FAUL TECHNICAL MEMORANDUM
FTM 70-1

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Searching and Editing Bibliographic Records Using the IBM Administrative Terminal System (ATS).

By Linda Webb
James Turner

May 1970

LI 002367

The work described herein was performed in the facilities of Syracuse University Library and the Technical Information Dissemination Bureau of the State University of New York at Buffalo.

CONTENTS

Introduction

Searching and editing bibliographic records using the IBM Administrative Terminal System (ATS) 1

FIGURE 1 Sample printout of Masfile-I records with field tags and manual addition of Syracuse University Library holdings statements 6

FIGURE 2 ATS Dialog Sequence: Method 2, using "LOCATE" and "PRINT WITH LINE NUMBER" commands 7

FIGURE 3 ATS Dialog Sequence: Method 3, using Searcher's Roulette technique 8

APPENDIX: Masfile-I Syracuse University Library project statistics, rate table and costs. 9

INTRODUCTION

The purpose of the FAUL Technical Memoranda (FTM) series is to disseminate quickly to librarians and information scientists the objectives, methods, procedures, analyses, conclusions, and recommendations relating to the performance of small projects in applied research. These projects may be imbedded in large long-term efforts, or they may have been undertaken to answer specific questions which bear upon the rational improvement of library procedures and services. Furthermore, such research may have been done wholly within a single library and by its own staff, or by the FAUL Central staff or committees, or by outside specialists under contract.

The material contained within these reports is used by appropriate FAUL staff in planning and decision-making, as well as for preserving an historical public record of the technical interests of the Five Associated University Libraries.

The material contained herein should not be reproduced in any form without written permission of the author. Publication in this series in no way denies an author from disseminating the information in other forms. Comments, corrections, and suggestions for improvement are most welcome.

These memoranda are routinely sent to the ERIC Clearinghouse on Library and Information Science for announcement in its publications and in Research in Education.

Coordinator of Library Systems
Five Associated University Libraries

SEARCHING AND EDITING BIBLIOGRAPHIC RECORDS USING THE IBM ADMINISTRATIVE TERMINAL SYSTEM (ATS)

Linda Webb and James Turner
Syracuse University Library
Research and Planning Group

BACKGROUND

In October 1968 the Five Associated University Libraries (FAUL)*/ committed resources to a series of experiments directed toward constructing a central file of bibliographic records in a cooperative environment. This series of experiments is called the Masfile Pilot Project. The first experiment, Masfile-I, was devoted to building a small file containing bibliographic records from the shelf lists of each library.

Since each member library uses the Library of Congress Classification Schedules to organize its collections, the FAUL Systems Committee decided to select the first five hundred shelf list cards in LC class Z116-Z1000.5, Book Trade and Library Science, from each library as the data base. A Masfile Input Group was formed composed of cataloging personnel from each library to advise the project and to implement the local arrangements. The shelf-list cards were collected and edited by a FAUL Central staff member, then forwarded to the Technical Information Dissemination Bureau at SUNY-Buffalo. The data were converted into machinable form on IBM2741 terminals connected to the Administrative Terminal System (ATS) running on an IBM 360/40. A printout containing approximately 1800 entries was produced on multiple-part forms and distributed to each of the five libraries.

Except in the case of Syracuse University, each member library compared the printout with local catalogs manually. Call numbers and quantities were added to citations which other member libraries submitted. These updated lists were mailed back to Buffalo and appropriate changes were made to the machine file by ATS terminals.1/

PROCEDURE

At Syracuse University Library, it was decided to modify Masfile records directly through terminals located at Syracuse and connected by dataphone to the ATS processor at Buffalo, a distance of about 170 miles. By this decision the adaptability of ATS to bibliographic searching, remote file access, and editing, could be tested in a practical environment.

In June 1969 therefore, the Technical Information Dissemination Bureau (TIDB) at SUNY-Buffalo transferred seven files containing the 1800 Masfile records to a storage area in the computer which a Syracuse terminal operator could access remotely. By this means, Syracuse Library staff could access the file whenever

---

*/ FAUL includes the libraries of SUNY-Binghamton, SUNY-Buffalo, Cornell University, University of Rochester, and Syracuse University.

1/ See Five Associated University Libraries *Masfile Pilot Project Final Report.* April 11, 1969 for additional details. (Available from the ERIC Document Reproduction Service, number ED 028 801)

the ATS was in operation at SUNY-Buffalo. The transfer process was simple: TIDB assigned the files to a Syracuse account number as stored messages, and the Syracuse terminal operator transferred the files/messages to Working and Permanent Storage areas within the ATS.

These files contained a total of 35,795 lines, or approximately 887 Permanent Storage Records (PSR). One PSR is equal to about 1500 characters in disc storage. Seven separate files (or ATS Documents) were needed because the ATS system limits the size of a Document to a maximum of 9,999 lines. The ATS Documents contained different quantities of lines since records of variable length were stored.

The staff at Syracuse proceeded to update Masfile-I records by adding the local call numbers of its holdings for books entered originally by other FAUL institutions. Because multiple call numbers for a single book were entered on a single line, 453 line entries were made containing a total of 481 call numbers. Excluding training time and delay in returning the updated files to TIDB, 23 hours and 48 minutes were needed for the project. An overall average of 3.00 minutes per entry, or 20.0 entries per hour was the measured rate for entry modification. The "per hour" time included time used to locate the position of the entry within the file and the typing and insertion of the new character strings. Also included in the figure was the time needed to sign-on to ATS, move an ATS Document from Permanent to Working Storage, send and receive messages to other terminals, store the updated file, and sign-off. Since these were all part of the standard operating procedure it seemed proper to include them. Besides, they occupied a very small percentage of the total time. More costly, time consuming, and annoying were the frequent telephone line interruptions which averaged over two per day; about one interrupt for each hour connected. The staff worked a total of twenty-eight days on the project and made sixty-five calls. Since standard telephone lines were used, a line interruption meant re-initiating the call at the regular three-minute rate. A direct line to TIDB, or a WATS line would have reduced the cost; having ATS available on the Syracuse campus would have eliminated the problem altogether.

THE EDITING SEQUENCE

The original printout of Masfile-I was used as well as an alphabetical listing of the records. Since ATS was used as the original data capturing technique, the order of the records in the file was operationally random; the original records were not sorted before conversion. The first step then was for searchers to check each entry using the alphabetical listing made from the ATS file against the Syracuse holdings. Call numbers were written on the alphabetical listing.

The second step was to manually transfer Syracuse call numbers from items on the alphabetical list to the same items as they appeared on the random Masfile-I list. (See Figure 1).

For comparison purposes three different methods of adding the call numbers to the machine file were used. The "print by line number" ATS command method was used first, resulting in a line-by-line printout of Masfile until the location of the next insertion was approached. This method was extremely slow and inefficient. However, the "locate" command was soon added to the ATS command set by SUNY-Buffalo, and the procedure was immediately changed to a second method. Using

"locate" and "print with line number" the process was speeded up, but it soon became evident that "locate" was also intolerably slow. In an ATS Document of 4860 lines an average of twenty seconds was needed to locate an entry. An ATS Document of 8000 lines took over a minute. The length of the string of characters to be located made no appreciable difference in the "locate" time. Since "locate" searches the entire ATS Document in working storage sequencially, its speed is related to the size of the Document. It was found that a Document of about 2000 lines was the practical limit for using the "locate"command. An efficient operator using "locate" and "print with line numbers" consumed an average of three minutes to locate the position of the entry, type the entry, and transmit it to the file. (See Figure 2).

SEARCHER'S ROULETTE

A third searching method was developed to increase the speed of interaction between the file and the operator. It was discovered that each citation in the list occupied an average of twenty lines. Since the line number of the last entry was known, it was possible to calculate the neighborhood of the first line number of the next entry, because each line in the ATS Document was numbered consecutively. The terminal operator counted the number of citations from the last edited entry through the next entry to be edited and multiplied it by twenty, adding this product to the line number of the last edited entry. For example: suppose the line number of the last entry of a Syracuse call number was 121; then three listings followed for which no Syracuse holdings existed and would therefore be skipped, but in the fourth listing a call number addition was needed. In this case multiplying twenty (lines per book listing) by four (the fourth entry away is to be modified) yields a product of eighty. Adding eighty to the line number of the last entry, 121, sums to 201. This sum is the approximate line number in the next bibliographic record to be edited. Now instead of using the "print with line number" command, which prints a minimum of five lines and is therefore slow, the single line print command is used which provides almost instantaneous response. Line 201 provides a "neighborhood" reference point which can be checked against the printout (see Figure 3). It was found that this calculation was usually within five lines of the next entry location.

Using this method, an efficient operator took an average of 1.5 minutes per entry, half the time needed using the "locate" and "print" method. A major drawback of the third method, dubbed "searcher's roulette", was the difficulty in finding an operator who could use it.

Searcher's roulette was devised by the library programmer, who had no difficulty using it and who did all the Syracuse Masfile updating. Attempts were made to train two other operators. Both were MT/ST operators. The first had a high school diploma and a few college credits in Business. She found the necessary calculations difficult and was unable to master the ATS commands.

The second operator also had a high school education in Business, but she found the technique challenging and was eager to learn it. The project was already completed before an attempt to train her could be arranged. However, it seemed likely that she could have mastered searcher's roulette in a reasonable time. An above-average intelligence coupled with a positive attitude toward machines seem to be the necessary requirements for ATS Masfile updating using searcher's roulette techniques.

OBSERVATIONS AND CONCLUSIONS

As a result of this project, the following observations and conclusions can be drawn concerning the use of ATS in this particular library environment:

1. ATS is an effective data collection device, but its searching capabilities are severely limited:

   a. Searching an ATS Document line-by-line is a time-consuming process requiring slow printout speeds.
   b. A search can operate only upon 9,999 lines of text at any one time.
   c. Working Storage limitations make it necessary to maintain a quantity of separate ATS Documents in Permanent Storage, each of which is brought in as needed and then returned to storage, so that a series of duplicate searches is necessary if the total file exceeds 9,999 lines.
   d. Storage costs as well as access limitations indicate that ATS work is practical only with relatively small and highly fluid files, unless it is used purely for data capture and text editing functions. Costs can be somewhat reduced by putting data into off-line storage when not needed and later returning it to the system, but this is not very practical in the present system configuration.
   e. To be accessible data must be maintained in a rigid ATS format. It must therefore be reformatted for other types of processing, but returned to ATS format if more ATS editing work is needed.
   f. Random searching cannot be done except in a very limited sense. Therefore, the searcher must have an "index" of the ATS Document by line number if searching is to be effective.
   g. Precise knowledge of the file structure and data location is mandatory for all practical purposes.

2. Telephone charges and line interruptions more than doubled the cost per entry in this project. Use of WATS or a direct line would result in substantial savings, although line interrupts would continue to be a problem, especially in WATS. Investigation of a direct line cost showed that at least two hours daily use of the terminal would be required for justification.

3. Long distance work of this kind presents no real problems, so long as the user stays informed about schedules, system changes, and the like. If data was to be processed other than in ATS (e.g., high-speed printer), transportation problems could arise.

4. The use of machine-oriented clerical personnel rather than a programmer would have reduced costs further.

5. Terminal printout should be restricted as much as possible, since the typewriter's effective speed is 148 words per minute. A rule of thumb of the project was "don't print any more than necessary."

6. Had the project been costed in terms of minimum time, not including experimentation, the cost would have been $806.34. Over $500.00 of the total cost went to experimentation. (See Appendix).

7. A revised "Locate" command now permits a terminal user to printout between line ranges e.g. 75-90. This command technique was not available during the experiment.

Syracuse University Library has not ruled out the use of ATS altogether, for it has its definite uses as a data capturing system and text editor when properly handled. It is being considered among other applications, for use in editing data captured with off-line devices. Editing of catalog data may be done manually from a listing produced by high speed printer from data captured on IBM MT/ST's, and corrections entered into the file via ATS, using line numbers for access. Corrected data would then go out of the ATS system for further processing by other computer software systems.

In general, this project was considered worthwhile as a demonstration of the potentials and limitations of terminal systems such as ATS. The splendid cooperation of the staff of the Technical Information Dissemination Bureau at SUNY-Buffalo and its director, Dr. Gerald Lazorick has been greatly appreciated by Syracuse.

001230

Computer assigned item numbers

90t None
90a 044412791632
10t Oehlerts, Donald D.
20t A study to determine the feasibility of establishing a cooperative technical processing program and direct transmission of interlibrary loans =. With a foreword by Le Moyne W. Anderson.
22t Study to determine the feasibility of establishing a cooperative technical processing program and direct transmission of interlibrary loans
30a Denver
30b Association of State Institutions of Higher Education in Colorado
30c 1962
40t x. 45 p. tables. 23 cm.
60b A report prepared for and submitted to the Association of State Institutions of Higher Education in Colorado.

Holdings statement for SUNY-Binghamton

60c Bibliography: p. 38-39.
70t Librarians -- Colorado.
70t Cataloging. Cooperative.
70t Inter-library loans.
72t Association of State Institutions of Higher Education in Colorado.
73t A cooperative technical processing program and direct transmission of inter-library loans.

ATS Document line number where new addition is to be made → line 1882

80b NBiSU-H/2732C603/ (1)

Z 732C 608/ (1)

Manually added holdings statement by Syracuse University Library

BUF 690203001230110HOLD 69/02/030000123001 35

001231

A new record begins here

90t None
90a 044502610703
10t Cessolis, Jacobus de.
20t The game of chesse=, by William Caxton. Reproduced in facsimile from a copy in the British museum.
(etc.)

These field tags do not appear in final printout of the book catalog

FIGURE 1: Sample printout of unedited Masfile-I record with field tags and manual addition of Syracuse University Library holdings statements

An * indicates lines typed by
the terminal operator

Commentary

Searcher is told which line number in the file the ATS routine is ready to operate on

Operator types a LOCATE (1) command to find out if the character string beginning NBiSu-H/etc. can be found. Then continues this strategy throughout the part of the dialog

A character string is located in the neighborhood of the line number which the searcher wants to edit

Print with line number command causes the computer to print out a bibliographic record preceded by field tags, followed by end-of-field indicators (±) and the line number in the file which contains the field

After checking with a paper printout of the file (which did not contain line numbers), the searcher locates the line after which the new line entry is inserted

```
NEXT NUMBER 4874
*1NBiSU-H/Z695.1C7s45
NOT FOUND
*1NBiSu-H/Z695.1C7s45
NOT FOUND
*1068215220272
NOT FOUND
*1Shariff, Waser, 1925
NOT FOUND
*1NBiSu-H/Z699S47
NOT FOUND
*1Sharp, Harold S ed.
NOT FOUND
*1NSyU/Z665S4
STRING 'NSyU/Z665S4' WAS FOUND ON LINES:
1241, 1329, 1405, 1424

*pn1250
SIGNAL WHEN READY
30b Scarecrow Press±                                    1250
30c 1964±                                               1251
40t 759 p. illus.  22 cm.±                              1252
60b Includes bibliographies.±                           1253
70t Information storage and retrieval -- Addresses/---/±
70t Libraries -- U. S. -- Addresses, essays, lectures.±
73t Title.±                                             1256
80b NBiSU-H/Z699S47/(1)±                                1257
80c NBuU/Z699S47/(2)±                                   1258
80d NIC/Z699S53/(1)±                                    1259
80e NRU/Z699S53r/(1)±                                   1260
92t 010.78±                                             1261
94t 64-11792±±                                          1262

90t None±                                               1264
90a 068515241583±                                       1265
10t Sharp, John R±                                      1266
20t Some fundamentals of information retrieval.±±       1267
22t Some fundamentals of information retrieval ±±       1268
30a New York±                                           1269

*80f NSyU/Z699S47/(1)±
*m1260
AFTER--  80e NRU/Z699S5  --MOVE--  80f NSyU/Z699S*y
NEXT NUMBER 4875
```

FIGURE 2: ATS Dialog Sequence: Method 2, using "LOCATE" and "PRINT WITH LINE NUMBER" commands

An * indicates lines typed by
the terminal operator

Commentary

Searcher is told which line number in the file the ATS routine is ready to operate on

Operator instructs ATS to go to line 1400 and print it on the terminal; the 73t---response is compared visually with the file printout and seen to be beyond the line where new data is to be added by the operator

After a sequence of approximations the correct part of the correct record is located and the new data inserted by the AFTER --- MOVE commands

```
NEXT NUMBER 2151
*1400
 73t Title.+
*1415
 60b Bibliography:  p. 131-133.+
*1405
 90t Z696.D5199+
*1402
 80c NBuU/Z696U55/ (4)+
*1401
 80b NBiSU-H/Z695U58/ (1)+
*1403
 94t 17-26005 rev++
*80f NSyU/025.  4U581/ca/ (1)+
*m1402
 AFTER-- 80c NBuU/Z606U --MOVE--  80f NSyU/025.4 *y
 NEXT NUMBER 2152
*1500
 90t Z695.95.U53 1956+
*1497
 80e NRU/A106U56m/ (1)+
*1494
 73t Title.+
*1495
 80b NBiSU-H/Z663.34M3/ (1)+
*80f NSyU/Z6621U58M2? (1)+
*m1497
 AFTER-- 80e NRU/A106U5 --MOVE-- 80f NSyU/Z6621 *y
 NEXT NUMBER 2153
 (etc.)
```

The "Y" indicates that the data is correct and the command should be executed

FIGURE 3: ATS Dialog Sequence; Method 3, using Searcher's Roulette technique

APPENDIX

MASFILE-I SYRACUSE UNIVERSITY LIBRARY PROJECT STATISTICS, RATE TABLE AND COSTS

I. STATISTICS

Masfile-I file size	1,828	records
	35,795	lines of text
Holdings additions made to the file by Syracuse	453	lines of text
lines added containing 1 call number	426	
lines added containing 2 call numbers	26	
lines added containing 3 call numbers	1	
Disk storage		
Quantity of lines of text per Permanent Storage Record (PSR)	40,825	
Quantity of PSR'S	887	
Quantity of characters per PSR	1,500	approx.
Total time used to manipulate Masfile by Syracuse		
in hours	23	hrs. 48 min.
in minutes	1,428	minutes
time used for housekeeping	81	
time used for editing	1,347	
Quantity of calls made from Syracuse to TIDB in Buffalo including re-establishing of unintentional disconnects	65	

II. RATE TABLE

ATS connect time	$2.00	per hour
(rate changed during project)	1.75	per hour
PSR storage	.005	per day per PSR
Long distance telephone from Syracuse to Buffalo	.70	for first 3 min.
	.20	for each add. min.

III. COSTS

	Base cost [1/]	Actual cost [2/]	Projected cost [3/]
cost per min.	$ .49	.725	.937
cost per entry, Method II ("Locate" plus "Print with Line Number")	$ 1.47	2.172	2.881
cost per entry, Method III (Searcher's Roulette)	.735	1.067	1.405
TOTAL COST	$704.09	1036.00	1339.00

1/ Base cost is what the project would have cost if everything worked smoothly and excess storage was not charged.
2/ Actual cost is the out-of-pocket expenses to perform the work of the project as described in the report.
3/ Projected cost would be project expenses if it were repeated under the revised cost structure imposed by the computer center part way through the project. The revised schedule reduced the connect time by about $1.00 but increased storage charges by $304. Moral: obtain rate guarantees for specific time periods before work begins.

PROJECTED COST COMPONENTS

Personnel	$ 225
Computer connect time	55
Disk storage charges	472
Long distance telephone	367
Data set rental	220
TOTAL	$1339